Theory of Error.

Let a community be divided into a number of classes, the proportions in the different classes being $z_1, z_2, z_3 \ldots$, so that $z_1 + z_2 + z_3 + \ldots = 1$. Suppose a random selection of n individuals to be made, the numbers drawn from the different classes being $n(z_1 + \epsilon_1)$, $n(z_2 + \epsilon_2)$, $n(z_3 + \epsilon_3) \ldots$ It is proved geometrically, with the aid of the binomial theorem, that the values of the errors $\epsilon_1, \epsilon_2, \epsilon_3 \ldots$ are normally distributed, and that the distributions are normally correlated. It follows that the values of any expression of the form $\Sigma \mathrm{A}\epsilon = \mathrm{A}_1\epsilon_1 + \mathrm{A}_2\epsilon_2 + \mathrm{A}_3\epsilon_3 + \ldots$ are normally distributed. The mean square of $\Sigma\mathrm{A}\epsilon$ is shown to be $\{\Sigma\mathrm{A}^2 z - (\Sigma\mathrm{A}z)^2\} \div n$, and the mean product of $\Sigma\mathrm{A}\epsilon$ and $\Sigma\mathrm{B}\epsilon$ to be $\{\Sigma\mathrm{AB}z - \Sigma\mathrm{A}z \,.\, \Sigma\mathrm{B}z\} \div n$. The applications are of two kinds:—

(1) The values of the probable errors in the determination of certain quantities are obtained, and, in particular, the probable errors in the mean, mean square of deviation, mean product of deviations, and divergence.

(2) Formulæ are obtained for testing particular hypotheses; *e.g.*, whether two distributions (of any kind) are independent; whether a distribution is normal; and whether two normal distributions are correlated.

"Mathematical Contributions to the Theory of Evolution. IV. On the probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation." By KARL PEARSON, M.A., F.R.S., and L. N. G. FILON, B.A., University College, London. Received October 18,—Read November 25, 1897.

(Abstract.)

1. This memoir starts with a general theorem, by which the probable errors made in calculating the constants of any frequency distribution may be determined. It is shown that these probable errors form a correlated system approximately following the normal law of frequency, whatever be the nature of the original frequency distribution, *i.e.*, whether it be skew or normal. The importance of this result for the theory of evolution is then drawn attention to. It is shown that any selection, whether of size, variation, or correlation, will in general involve a modification not only of the size, but the variation and correlation of the whole complex of correlated organs. The subject of directed selection, of which this random selection is only a special case, is reserved for another memoir, nearly completed.

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to *Proceedings of the Royal Society of London.* www.jstor.org

2. Normal correlation is first dealt with. It is shown that if $\lambda = 0{\cdot}67449$ and n be the number of observations:

Probable error of a coefficient r_{12} of correlation $= \lambda \dfrac{1 - r_{12}^2}{\sqrt{n}}$

Probable error of r_{12}, for variations with definite values,*

$$= \lambda \frac{1 - r_{12}^2}{\sqrt{n}} \frac{1}{\sqrt{1 + r_{12}^2}}.$$

Correlation of errors in two standard deviations $= r_{12}^2$.

Correlation of errors in a standard deviation and a correlation coefficient $= r_{12}/\sqrt{2}$.

Probable error of a regression coefficient for two organs

$$= \lambda \frac{\sigma_1}{\sigma_2} \frac{\sqrt{1 - r_{12}^2}}{\sqrt{n}},$$

where σ_1 and σ_2 are the two standard deviations.

Probable error of a regression coefficient for three organs

$$= \frac{\lambda}{\sqrt{n}} \frac{\sigma_1}{\sigma_2} \frac{\sqrt{1 - r_{23}^2 - r_{13}^2 - r_{12}^2 + 2 r_{12} r_{23} r_{13}}}{1 - r_{23}^2}.$$

Correlation between the errors in two correlation coefficients, *i.e.*, r_{12} and r_{13}

$$= r_{23} - \frac{r_{12} r_{13} \left(1 - r_{23}^2 - r_{13}^2 - r_{12}^2 + 2 r_{12} r_{23} r_{13}\right)}{2\left(1 - r_{13}^2\right)\left(1 - r_{12}^2\right)}.$$

Correlation between the errors in two correlation coefficients, *i.e.*, r_{12} and r_{34}

$$= \frac{\left\{\begin{array}{l} \quad (r_{13} - r_{12} r_{23})(r_{24} - r_{23} r_{34}) + (r_{14} - r_{34} r_{13})(r_{23} - r_{12} r_{13}) \\ + (r_{13} - r_{14} r_{34})(r_{24} - r_{12} r_{14}) + (r_{14} - r_{12} r_{24})(r_{23} - r_{24} r_{34}) \end{array}\right\}}{2\left(1 - r_{12}^2\right)\left(1 - r_{34}^2\right)}.$$

3. Skew variation is next dealt with. First the probable errors and correlations of the errors of the constants of the curve

$$y = y_1 \left(1 + \frac{\gamma x}{p + 1}\right)^p e^{-\gamma x}$$

---

* This value for the "array" was erroneously given as that for the absolute value of r_{12} in 'Phil. Trans.,' A, vol. 189, p. 345, but the statement was corrected in 'Roy. Soc. Proc.,' vol. 61, p. 350.

are calculated. The formulæ obtained have been cited and used in a memoir by Karl Pearson and Miss Alice Lee, "On the Distribution of Frequency (Variation and Correlation) of the Barometric Height at divers Stations."* We may note the following results :—

Probable error of mean $= \lambda\sigma/\sqrt{n}$.

Probable error of standard deviation σ

$$= \frac{\lambda\sigma}{\sqrt{2n}}\left(1+\tfrac{1}{2}\frac{1}{(p+1)^2 \mathrm{S}}\right)^{\frac{1}{2}},$$

where $\mathrm{S} = \frac{\mathrm{B}_1}{p} - \frac{\mathrm{B}_3}{p^3} + \frac{\mathrm{B}_5}{p^5} - \&c.$, and B_1, B_3, $\mathrm{B}_5 \ldots$ are Bernoulli's numbers.

Probable error of skewness S_k

$$= \lambda\sqrt{\frac{3}{2n}}\,\frac{1}{\sqrt{1+3\,(\mathrm{S}_k)^2}}, \text{ approximately.}$$

Correlation of errors in mean and standard deviation

$$= -\sqrt{\frac{2}{p+1}}\,\frac{1}{\sqrt{1+\frac{1}{2}\frac{1}{(p+1)^2\mathrm{S}}}}.$$

Correlation of errors in mean and skewness $= 0$.

Correlation of errors in skewness and standard deviation

$$= \frac{1}{\{1+2\,(p+1)^2\mathrm{S}\}^{\frac{1}{2}}}.$$

Thus a random selection of size differing in its mean value from the population mean gives in all probability an alteration in variability, but none in skewness. Increased size means decreased variability. A random selection altering variability alters also both size and skewness of distribution.

Probable error of modal frequency y_0

$$= \lambda\frac{y_0}{\sqrt{2n}}\left(1+\frac{1}{12p}\right)^{\frac{1}{2}}, \text{ approximately,}$$

which is shown to be always less than the probable error of the mean.

The correlation of errors in the moments and their probable errors up to the fourth are also worked out.

4. As a limiting case the probable errors of the skewness, the moments up to the fourth, and of the distance between mean and

---

* About to be published in 'Phil. Trans.'

mode in the case of a random selection from a normal distribution are worked out.

5. Analogous results are next obtained for the skew frequency curves:

$$y = y_1\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2}$$

and

$$y = y_1\frac{1}{\{1+(x/a)^2\}^m}e^{-\nu\tan^{-1}\frac{x}{a}}.$$

It is shown that in these cases the mean size, variation, modal frequency, and skewness are in general all such that their errors are correlated. Hence any selection of size modifies both the variation and skewness of the distribution. This is of considerable importance for the theory of evolution, as in most cases of zoological frequency the distribution is of the second, or tangent curve, type. Hence random selection of size tends to modify not only variation but skewness of distribution. The results are too long to be cited, and their application to special cases involves somewhat lengthy, but not complex arithmetic, which in practical cases we have found much shortened by the use of the "Brunsviga" calculator.

"Further Observations upon the Comparative Physiology of the Suprarenal Capsules." By SWALE VINCENT, M.B. (Lond.), British Medical Association Research Scholar.* Communicated by E. A. SCHÄFER, F.R.S. Received November 2,—Read November 25, 1897.

(From the Physiological Laboratory, University College, London.)

In previous communications† I have given experimental evidence in favour of the view that the paired suprarenal bodies and the interrenal gland of Elasmobranch fishes correspond respectively to the medulla and cortex of the suprarenal capsules of the higher Vertebrata. I have further stated, as the result of numerous experiments, that the medullary portion of the suprarenal appears to be absent in Teleosts, the suprarenal bodies in this order of fishes consisting solely of cortex.

Since performing the above series of experiments my attention has been devoted to the general physiological effects of extracts

---

* The expenses involved in this research have been defrayed by a grant from the Government Grant Committee of the Royal Society.

† 'Physiol. Soc. Proc.,' March 20, 1897; 'Roy. Soc. Proc.,' vol. 61, p. 64, 1897; 'Anat. Anz.,' vol. 13, Nos. 1 and 2, 1897.